ABSTRACT

A method and system for determining similarity between a first event set, the first event set including a first plurality of event types, and a second event set, the second event set including a second plurality of event types, is provided. Observed events are randomly mapped to a multidimensional vector-Q and query events are mapped to a multidimensional query vector-q. Comparison of the vectors for a predetermined similarity according to: $||Q-q|| \leq SV$, where SV = a predetermined similarity value determines similarity.